A Relatively High-Resolution Reading Aid for the Blind

JAMES C. BLISS

Abstract—**Many direct-translation reading aids for the blind have been built in the past, employing an auditory output consisting of a combination of tones indicating the black regions in a vertical slice through a letter space, or a tactile output consisting of a raised or vibrating image of the letter shapes. Maximum reading rates obtained by the majority of subjects with these reading aids have been less than 10 correct words per minute, and the cause of this limitation has not been well understood. By analyzing the spatial spectral content of letter patterns, we show here that most of these reading aids have violated the well-known sampling theorem. It is suggested that this may be a significant factor in the observed reading-rate limitation.**

The design of a reading aid is described based on the conclusions from this analysis of the sampling process, and on recent results from tactile research. With this reading aid a hand-held probe images a vertical section of a letter space onto a 24 × 6 array of photosensors, and the probe is manually moved horizontally across the line of print. The signal from each photosensor controls a tactile stimulator in a corresponding array of 144 stimulators, which are placed on a single finger.

In preliminary reading tests with this device, four subjects have all read at rates greater than 10 correct words per minute, and two of the subjects have read at rates greater than 20 correct words per minute. Possibilities for obtaining still higher reading rates are discussed.

INTRODUCTION

IN 1914 Fournier d'Albe [1] demonstrated a direct-translation reading aid for the blind with which printed material was moved horizontally across an image of a vertical row of illuminated dots. The light in each dot was chopped at a different rate, and the total light, reflected from all the dots on the page, was sensed by a single selenium bridge. The output signal from the selenium bridge was composed of the modulation frequencies in the reflected light, and these frequencies were arranged so that tones of low pitch represented low parts of a letter space, and higher tones represented higher parts of the letter space. The user was expected to recognize the letters by listening for the missing chords and tones as each letter was scanned.

Since that time many conceptually similar direct-translation reading aids with auditory output have been suggested and built [2]–[4]. Analogous reading aids have also been built that produce tactile images embossed on hard copy [5], or displayed on matrices of tactile stimulators [6]. However, none of these devices has become widely accepted or available. Several common properties that these reading aids share are 1) maximum reading rates obtained by the majority of the subjects have been less than 10 correct words per minute, 2) fewer than 12 photocells typically span the vertical dimension of a letterspace, and 3) the output signals from a single scan of the probe are not always sufficient for exact letter identification. In this paper evidence will be given indicating that a major cause of the 10 words per minute limitation on reading rate has been due to poor resolution, or to too few photocells in the device, and we will illustrate the importance of accurate single-letter identification in obtaining satisfactory rates with such devices.

Previously, we described a basic design for a direct-translation reading aid with a tactile output [6]. In that design an area slightly smaller than a letter space is imaged on an array of photosensors. The signal from each photosensor controls a tactile stimulator in a corresponding array of tactile stimulators. A unique feature of the design is the use of piezoelectric reeds for tactile stimulators, which are arranged to drive a dense array of vibrating pins efficiently. In this previous paper we also described some performance tests with a computer simulation of the device. The computer tests indicated that reading rates of 30 correct words per minute were possible by tactually sensing perfectly formed and registered letters.

In the present paper we will also describe the construction of a completely self-contained reading aid based on the previously described design and computer simulation tests. Subsequent tests with this reading aid have supported the validity of the design and the computer simulation experiments.

DESIGN

Resolution Considerations

In the design of a reading aid a primary consideration is the spatial resolution required of the images. Since high-quality optical images are easily obtained, this consideration mainly affects the choice of the number of photosensors, and the corresponding number of tactile stimulators to be used. Of course, a single movable photosensor could in principle permit a person, by scanning, to obtain eventually enough information for letter identification. However, if the scan is manual, reading would be extremely slow, and if it is automatic, perceptual considerations limit the scan rate so that a slow reading rate would also result. Therefore, some parallel channels are necessary. In this paper we will consider the case of completely parallel input for the vertical dimension and a single horizontal scan. We will also consider parallel input in the horizontal dimension but, since the scan is horizontal, these channels will pro-

Manuscript received October 15, 1968; revised November 14, 1968. This work was primarily supported by the Social Rehabilitation Service under Grant VRA-RD-2475-S-67, and, in part, by the Office of Education under Grant OEG-0-8-07112-2995(932), both with Stanford University.

The author is with Stanford University, Stanford, Calif., and Stanford Research Institute, Menlo Park, Calif.

vide mainly redundant information and their number will depend on perceptual considerations.

A lower bound on the number of photosensors required in a two-dimensional matrix spanning a letter space is given by $\log_2 N$ where N is the number of patterns to be distinguished. However, such a matrix would severely restrict the input pattern shapes producing a unique output, and the noise and registration tolerances of these patterns. A more realistic estimate is based on the well-known sampling theorem, which permits estimation of the number of photosensors required to reproduce the letter patterns faithfully. While it might be argued that the redundancy of English text makes it possible to read with poorer resolution than this, there are certainly many situations in which the contextual constraints are few or ambiguous, or the printing is poor, or some new symbol or type font is used, so that absolute letter identification is required for the reader to obtain the correct meaning. In addition, text in which each letter can be easily identified correctly can be read more rapidly than text composed of poorly resolved letters. Therefore, we will take the point of view that the photosensors sample the image plane, and that enough samples must be taken to reproduce each letter pattern faithfully.

Since the patterns move horizontally across the photosensor matrix, we need only to determine the sampling requirements in the vertical dimension. As shown in Fig. 1, this one-dimensional system can be considered to be equivalent to a combination of the transformations performed by an array of photosensors of infinite extent sampling the image plane and by an aperture. The corresponding spectra for each of these functions are also shown. Note that the further apart the photosensors are, the closer together the individual components of the output spectrum become. If the photosensors are too far apart, the individual spectra overlap, resulting in lost information. The maximum photosensor spacing d to prevent this overlap of the spectra can be determined from the spatial bandwidth of the input patterns. With a large aperture the requirement is that $1/d$ must be greater than $2s$, where s is the maximum significant spatial frequency of the input pattern.

An upper bound on the spatial frequency bandwidth of the input patterns s can be obtained by assuming that the information in these patterns necessary for identification is within the limitations of visual acuity. Actually, this is the only constraint on the evolution of the design of alphabetic shapes. Acceptance of the resolution implied by this constraint means that the reading aid should equal visual acuity over the height of a letter space. Campbell *et al.* [7] has measured visual contrast sensitivity[1] for horizontal gratings of various spatial frequencies, and these data are given in Fig. 2.

[1] By visual contrast sensitivity, is meant the reciprocal of the percent modulation at which a subject can detect, with high probability, the difference between the modulated field and a uniform field with the same average illumination.

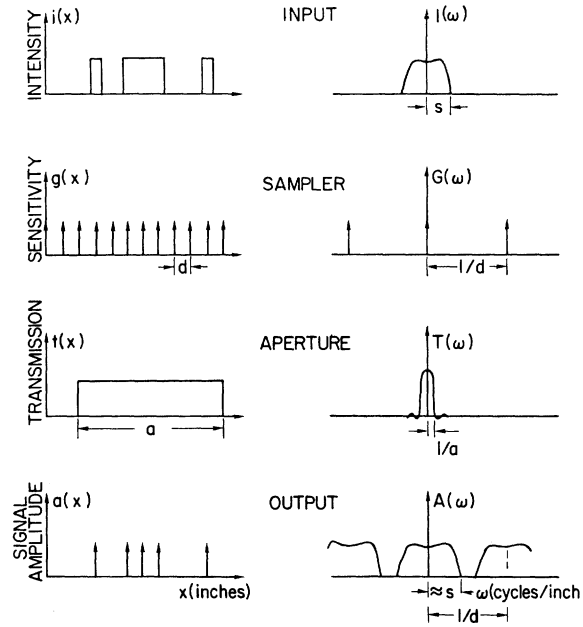

Fig. 1. Implications of image sampling with a phototransistor array. The image light distribution $i(x)$ is presumed to have a Fourier transform $I(\omega)$. Uniform sampling of this image with the function $g(x)$, limited by the aperture $t(x)$, results in the output $a(x)$. The corresponding spectral operations are $G(\omega)$ convolved with $T(\omega)$ and convolved with $I(\omega)$, which gives $A(\omega)$. For large a, $1/d$ must be greater than $2s$ to prevent loss of information.

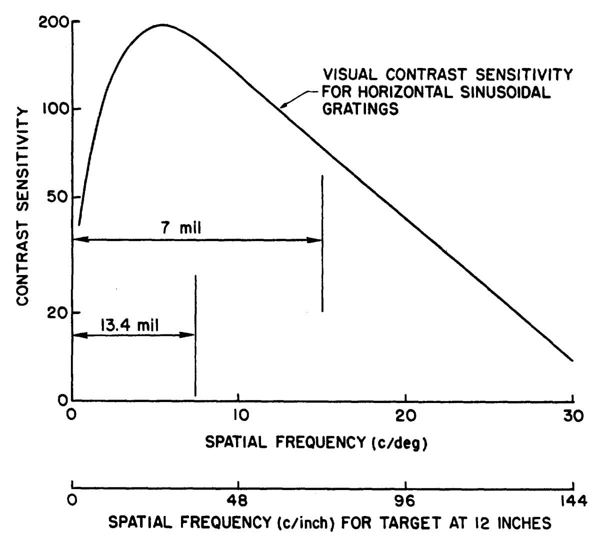

Fig. 2. Comparisons among visual contrast sensitivity, the spatial frequency bandwidth of pica type, and the bandpass of two sampling rates. Contrast sensitivity data are from [7]. For a reading distance of 12 inches, visual sensitivity is maximum at 24 cycles/in. Pica type has a bandwidth of about 75 cycles/in and a photosensor spacing less than 7 mils (referred to the page) is required to pass this bandwidth. The 7-mil and 13.4-mil dimensions indicate the bandwidths obtained with these photosensor sampling rates.

Related to these considerations, Harris [8] illustrates that a spatial bandwidth of more than five cycles across the vertical dimension of a block letter *s* is required to distinguish it visually from the block numeral 5. If the *s* were lower case, an adequate system would need a bandwidth of about 12 cycles across the full vertical dimension of a letterspace. Thus, to obtain this bandwidth, roughly 24 samples across the letter space would be required.

Another estimate of the spatial frequency bandwidth of alphabetic shapes, also shown in Fig. 2, is derived from direct measurements on pica type. These measurements indicate that most of the spectral energy of pica type is within a bandwidth of 75 cycles/in. Therefore, these various considerations all independently indicate that most of the information important for recognition, as well as most of the spatial spectral energy is commonly used type fonts, is contained within a bandwidth of about 75 cycles/in.

From these estimates it appears reasonable that a reading aid should have a resolution bandwidth of at least 75 cycles/in, which means an equivalent photosensor sampling rate (referred to the page) of 150 samples/in, or a photosensor spacing of about 7 mils. Since the distance from the top of the highest letter to the bottom of the lowest letter is about 160 mils for pica type, at least 24 photosensors are needed in the vertical dimension. This number agrees well with the number of vertical samples used in practice in automatic character-recognition equipment [9] and in facsimile equipment. However, most previous reading aids for the blind have used fewer than half this number of photosensors, resulting in an equivalent photosensor spacing greater than 13.4 mils (referred to the page). The bandwidth resulting from this photosensor spacing is also shown in Fig. 2, compared with visual acuity and the bandwidth of pica type.

The above analysis assumes that the photosensor samples are delta functions, that is, their individual apertures have zero area. However, photosensor sensitivity is proportional to photosensor area, so that a compromise must be made. If the photosensor signals are quantized into two levels, corresponding to black and white, then the smallest photosensor that permits adequate sensitivity should be used, since larger photosensors reduce the contrast between the two levels. The effect of photosensor area on contrast has been analyzed by Kazimerczak [10], with the conclusion that the photosensor aperture should be equal to or smaller than the linewidth of the letters in order to obtain an optimum contrast. Pica type commonly has 5-mil lines, so that the photosensor aperture should be slightly less than the spacing between photosensors.

If fewer than 12 photosensors are used to scan the 160 mils of a letter space vertically (as the case in most previous reading aids), and each photosensor has a field of view less than 7 mils, then it is possible for a 5-mil horizontal line to be completely undetected, its image falling into the insensitive space between photosensors. This condition alone can prevent unambiguous letter identification. For example, the lower case *e* could be indistinguishable from the lower case *c*.

Perceptual Considerations

Given a reading aid with at least 24 parallel channels, there is the psychological question of whether its output would be within the capabilities of human tactile perception. In an experiment with 24 well-spaced tactile stimulators, we measured the information conveyed by combinations of stimulators simultaneously activated for a brief period [11]. This experiment indicated that well-trained subjects had essentially all of the information in these tactile patterns available to them for approximately 1 second after termination of the stimuli. Thus, not only are 24 parallel channels necessary to convey adequately the printed information without vertical scanning, but they are also within the perceptual capabilities of the human tactual system.

The number of photosensors required in the horizontal dimension depends on psychological considerations, since all the information passes by a single vertical column of photosensors. In our previously reported experiments with a computer stimulation of the reading aid [6], we varied the number of vertical columns between 1 and 4. The results indicated that reading performance improved markedly if more than one vertical column was used, and little improvement resulted as the number of columns was increased from three to four. However, if the field of view could be made big enough for word-at-a-time perception rather than letter-at-a-time perception, another increase in reading performance might be obtained, as indicated by Troxel's [12] experiments.

The effect of different spatial sampling rates is illustrated in Fig. 3, which shows a letter moving across a 12×4 array compared to the letter moving across a 24×6 array.

Equipment Design

For scanning flexibility and for simplicity, we chose to build a hand-held optical probe, rather than a mechanical scanner. This choice implies that small size and weight are desirable so that the probe can be easily and rapidly manipulated. This requirement in turn means that a small image magnification is desirable. We chose the FPM 100 phototransistor for the light transducer because of its high sensitivity and small size. This phototransistor is encased in a can 80 mils in diameter, so that a magnification of at least 12 is required to form an image of a letter space on a 24×6 array of these phototransistors. Assuming a page-to-lens distance of about 1 inch (so that the working distance is large enough to permit illumination of the page), a magnification of 12 implies a 12-inch lens-to-phototransistor array distance. A dimension of this magnitude would result in an unwieldy probe. Moreover, because of the phototransistor encapsulation, a great deal of light would be lost between the photosensitive areas.

These problems can be overcome if most of the magnification is obtained by using an array of flexible fiber optics, one fiber per channel. Plastic fiber optics for this purpose are convenient and inexpensive. Thus in our design we chose a lens system to form a $1.6 \times$ magnified image on a 24×6 array of 10-mil diameter fibers. The fibers are bundled into a flexible cable, which transmits

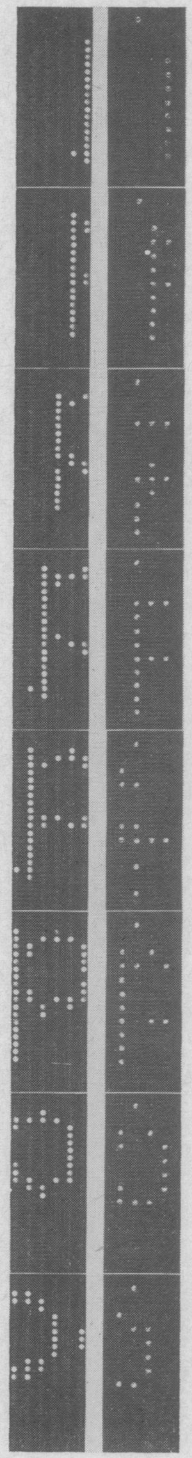

Fig. 3. Letter patterns resulting from image sampling with 12×4 and with 24×6 phototransistor arrays. As the optical probe is moved from left to right across the print (with arbitrary vertical registration), the patterns on the 24×6 array are recognizable a much greater percentage of the time than those on the 12×4 array.

the image to the phototransistors and electronics, each fiber being connected to a corresponding phototransistor. This arrangement enabled us to build a probe small enough to be moved sufficiently close to the binding of most books to scan the complete line of print.

In addition, a mechanical tracking aid was built to make it easier to scan along the line of print. This device limits the probe to horizontal movements, unless a lever is pushed that releases the vertical lock. Fig. 4 shows the complete system including the tactile stimulator array, the mechanical tracking aid and optical probe, the electronics, and a light display to indicate

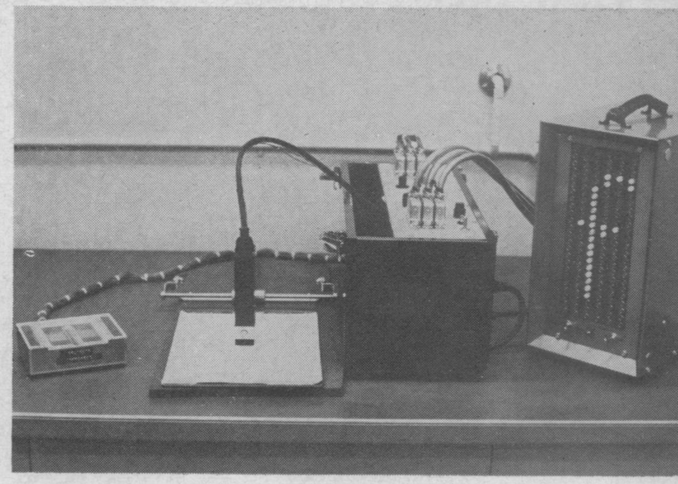

Fig. 4. The complete self-contained reading aid. From left to right are the tactile display, the mechanical tracking aid and the optical probe, the electronics, and a light display indicating activated tactile stimulators. The optical probe is connected to the electronics with a flexible bundle of fiber optics.

which tactile stimulators are activated. The light display is used by the experimenter in observing the reading performance of the subjects.

The electronic system consists of some control circuitry and 144 identical channels, each including a phototransistor, flip–flop, and tactile-stimulator transistor switch. The phototransistors are operated in the charge storage mode [13] with a storage time of 20 ms. Thus the output information is updated every 20 ms, well within the perceptual time of the user. Fig. 5 gives a schematic diagram of each channel in the electronic system.

Tactile Stimulator Array

The choice of body location for the tactile display depends on tactile sensitivity, resolution, area of skin available, and convenience. Geldard [14, pp. 308–311] argues that stimulators should be distributed as widely as possible over the body surface. However, when his data with widely distributed stimulators are compared to our finger stimulation data [11, p. 98] little, if any, difference can be noted. Spiridon [15] directly compared widely distributed stimulation with finger stimulation, and also found no significant difference in performance. Moreover, we have some evidence [16] that temporal resolution is better with closely spaced stimulators than with widely distributed stimulators, so that it may be important to have closely spaced stimulators for accurate perception of *moving* spatial patterns.

Weinstein [17] investigated pressure sensitivity, point localization, and two-point discrimination at 20 body sites. The most important of these measures for spatial pattern perception is two-point discrimination. (However, none of these measures directly measure pattern-perception capabilities.) The middle finger and index finger were the best, respectively, in two-point discrimination, and the index finger was best, in point localization, of all the sites tested. Perhaps a useful figure of

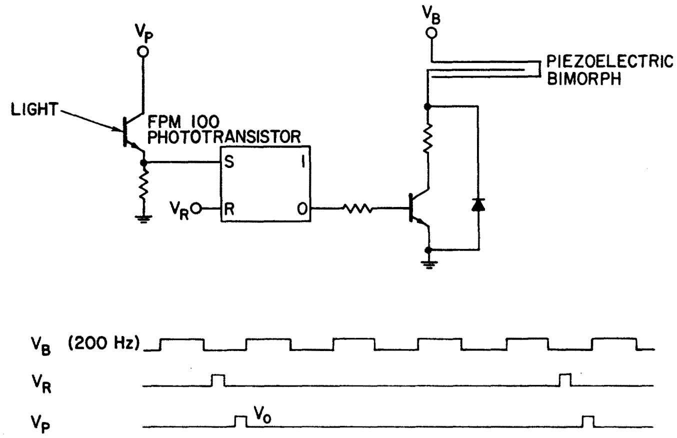

Fig. 5. The electronic system of the reading aid. The phototransistor base-to-collector depletion capacitance is charged to approximately V_o by the pulse on its collector V_p. The emitter junction is then back-biased. The light on the phototransistor discharges the depletion capacitance by an amount approximately proportional to the intensity of the light and the length of time between pulses. The phototransistor output pulse sets the flip-flop if the integral of the light over the preceding interval was greater than a threshold value. The set-reset flip-flop activates the piezoelectric stimulator through a transistor switch.

merit to consider in choosing a body site for a tactile display of spatial patterns is the ratio of skin area available to the square of the two-point threshold. On this basis also, the smooth skin of the fingers and palm were superior. For example, this ratio is several times greater for the hands than for the back.

The two-point discrimination threshold determined was approximately 100 mils on the fingertip. However, some caution should be used in taking this number too seriously because Mukherjee [18] has shown that, with practice, the two-point threshold decreases markedly, and, as long as the wavelength of the highest spatial frequency of the patterns is greater than the two-point threshold, the stimulator spacing can be less than this threshold.

The density of nerve endings in the fingertip is several hundred/mm² [19], much more than adequate to explain the psychologically determined two-point threshold. However, these endings are not all independent, and any one nerve fiber may have a receptive field that is several dermal ridges wide [20]. Taking into account overlapping of receptive fields, the limitation measured by the two-point threshold does not seem to result from the peripheral nervous system, so that it is not surprising that discrimination improves with practice. Thus, it appears that stimulators spaced slightly closer than the nominal two-point discrimination threshold would not exceed tactile perceptual capabilities.

These various considerations all suggest that a two-dimensional array of closely spaced stimulators on a fingertip would take maximum advantage of human characteristics and capabilities for tactile pattern perception. The next question to be considered is the form of stimulation at each point in the array.

Temperature, electricity, pressure, and vibration could be physical stimuli in such an array. Of these, we chose vibration because of the convenience and simplicity of the piezoelectric bimorph as a stimulator, and because an acceptable sensation is obtained with good two-point discrimination. Also, these stimulators require less power for adequate sensation than any we have found, and they can be closely packed relatively easily.

Several investigators [21], [22] have found evidence in glabrous skin of man and monkey for two different peripheral neural systems responsible for the sensation of flutter or vibration. One of these systems of receptors appears to be more sensitive at low frequencies (20–40 Hz) and to have better spatial resolution than the other system. The evidence also indicates that the low-temporal-frequency high-spatial-resolution system consists of dermal ridge receptors, while the other system lies in deep tissue and probably terminates in Pacinian corpuscles. The low-spatial-resolution system is extremely sensitive at optimal frequencies of 150–250 Hz. This system is thought to produce sensations described as vibratory hum deep within the hand, which is so spread that accurate localization is difficult.

On this basis one might expect 20–40 Hz to be the optimum frequency for stimulator vibration. However, letter shapes moving past a fingertip, produce a transient (rather than steady-state) condition at any point on the skin. Lindblom's [22], [23] data from the glabrous skin of monkeys were taken under similar transient conditions, and data indicate that both neural systems would be stimulated by at least the first few pulses in any pulse train with a frequency up to 100–200 pulses/s. The hypothesis that both neural systems respond over a wide range of stimulator frequencies in a tactile reading situation is supported by the following experiment, which C. H. Rogers and the author have performed.

With a computer simulation of the reading aid and a 12×8 array of airjet stimulators, we varied the frequency of the air pulses from 24–160 pulses/s, keeping a constant pulse width of 1 ms. The stimulus patterns consisted of a "moving belt" of random letters analogous to certain electric light display signs. The "belt" velocities used correspond to 12 and 30 words/min. Subjects were trained by requiring them to name single moving letters, and immediate feedback regarding the correct answer was given. After single-letter identification accuracy reached 80 percent, testing was initiated. The test stimuli consisted of random-length strings of from two to eight random letters. After each string moved across the airjet array, all airjets would be turned off, and the subject would attempt to report the next to the last letter displayed. This paradigm was devised to avoid subject guessing based on context, to simulate actual letter spacing, and yet not to overtax short-term memory capabilities in responding.

Fig. 6 gives the percentage of letters correctly reported as a function of the frequency of the airjet pulses and the "belt" velocity. The upward trend of accuracy as the frequency is increased is significant at both the 12-word/min and 30-word/min rates. Thus,

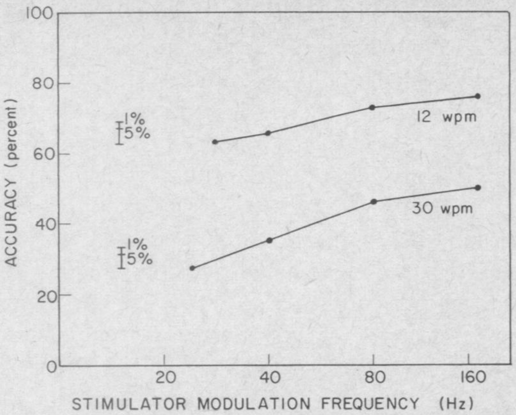

Fig. 6. The effect of tactile stimulator modulation frequency on letter recognition accuracy. The random letters were moved across a 12×8 airjet array by a computer at a rate corresponding to either 12 words/min or 30 words/min. Data are averaged over three subjects, and each point is the average of 300 trials per subject or 900 trials total.

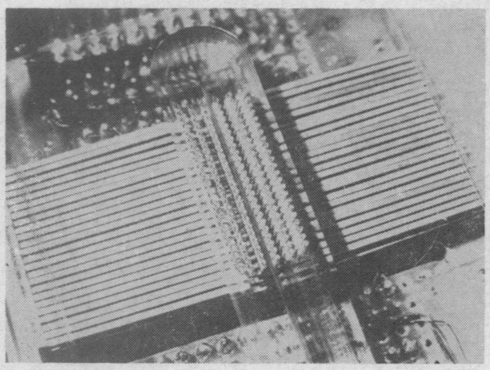

Fig. 7. Tactile stimulation array. The 24×6 array of tactile stimulators fits on a single fingertip. The stimulator pins are spaced 50 mils apart along the finger and 100 mils apart across the finger. The perforated surface is curved to fit the finger.

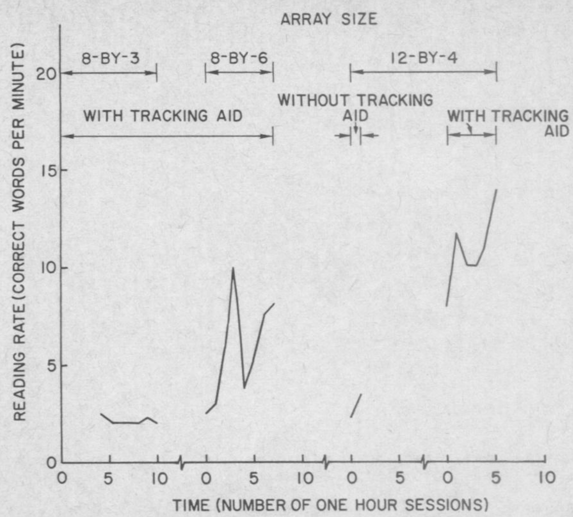

Fig. 8. Reading performance of one subject under a variety of conditions. These reading tests were made over a period of two years with various modifications of two reading aids. The printed material was moved across the optical pickup with the tracking aid used in the first 17 sessions. With the tracking aid used in the last five sessions the optical pickup probe was moved across the printed material.

higher stimulator modulation frequencies appear to result in increased performance.

The design of the tactile stimulator array shown in Fig. 7 is consistent with these results and considerations. The piezoelectric reeds in this array are driven near their resonant frequency, which is 200 Hz. The 24×6 array of 144 stimulators fits on a single finger, spanning the distal and a portion of the middle phalanges. The unloaded reeds deflect about ± 30 mils from their rest position, and loaded with a fingertip, with contact occurring only over a small portion of a cycle, the deflection may be reduced to roughly ± 5 mils. The stimulator pins that contact the skin are 10 mils in diameter, and the top of the pins and the perforated surface are curved to fit the finger.

Evaluation

The ultimate effectiveness of a reading aid based on the principles described above depends on how difficult it is to use, how much training is required to learn to read with it, and what reading performance can be obtained with it. While extensive testing has not yet been performed with this reading aid, we do have some preliminary results on several subjects.

The first subject, CL, was a high school senior girl, blind from an early age, who was also a subject in our computer simulation experiments, and so was asked sporadically to give her impression of various earlier versions of this reading aid. In many of these latter sessions, reading rates were measured, and these data are summarized in Fig. 8. By the time these data were taken, this subject knew the upper and lower case letter shapes, and was proficient with recognition of letter shapes displayed on an array of vibrating tactile stimulators. Therefore, the general upward trend of the points is most likely due to improvements in the apparatus rather than to learning. Note that the reading rate generally increased with increases in array size and increases in the number of vertical channels, and that significantly higher reading rates are obtained with a tracking aid.

The other three blind subjects, LS, RS, and JK, had no previous experience with a reading aid. With two of these subjects the lower case letter shapes had to be learned, and all three subjects varied considerably in their ability to spell and in their facility with written language. LS was a college sophomore, RS a computer programmer, and JK a recent college graduate.

To directly verify the resolution requirements described above, a legibility experiment was performed with the reading aid. Random strings of upper case letters and numbers, and lower case letters, were printed in four sizes of Mid-Century typescript. (This printed material was identical to that used by Arps *et al.* [24].) Each letter and number was manually scanned with the reading aid by two sighted and two blind subjects (CL and RS). All four subjects were instructed to take as much time as they needed to make each identification. The sighted subjects made their identifications by observing the light display, and the blind subjects used the tactile array. The performance accuracy of each group of subjects is shown in Fig. 9. Legibility in the 92–98 percent range was obtained at the letter-space height that the reading aid was designed for (i.e., 160 mils or 24 samples across the height of the letter space). Since the size of letters on the light display in no way taxed visual acuity, the sighted subjects' performance primarily indicated sampling rate influences on legibility. For sampling rates less than the design value, legibility dropped rapidly. For example, with the 71-mil letter-space height (equivalent to about 10 photosensors across the height of the letter space), a visual lower case letter legibility of 81 percent was obtained. Tactile performance was significantly worse than this, probably because of the smaller size of these letters, as well as the poorer resolution. Although some reading is possible with 81 percent legibility, it is slower, less accurate, and, generally, unsatisfactory.

The performance of all four blind subjects with the 24 × 6 reading aid is shown in Fig. 10. The material read for these tests was from *Reading Skill Builder* (stories and articles adapted from the *Reader's Digest*) 1958 and 1959, published by Reader's Digest Services, Inc.

In these tests, unlike our experience with lower resolution reading aids, subjects could correctly identify letters at a much earlier stage of training. No help was given by the experimenter when a difficult word was encountered. The rates shown are, therefore, realistic measures of completely independent reading at this low level of training. Since the three new subjects achieved reading rates of at least 10 words/min in less than 20 hours of using the device, we believe that this reading aid has considerable promise. While two subjects at this writing can read at over 20 words/min, and the reading rates of all four subjects are still increasing, an ultimate reading rate has not yet been established. In fact, maximum reading rates will probably not be attained until the reading aid is used in everyday life.

The rate of learning to read with this aid appears to be significantly faster than with previous lower resolution auditory and tactile direct-translation reading aids. Differences in training procedures mean that our learning curves are not strictly comparable with those from other reading aids. For example, considerable time was spent in the Battelle program learning individual letters and words before actual reading was begun. Our subjects began reading on the first day. However, it appears from the reported literature that the Battelle students required 65 hours to obtain a reading rate of about nine words/min, and that the subject with the Mauch Visotactor required 130 hours to obtain a reading rate of about eight words/min. The more rapid progress of our subjects may be because of the simplicity of the direct correspondence between the printed letter shape and the two-dimensional tactile pattern, and because of the relatively greater resolution of our reading aid.

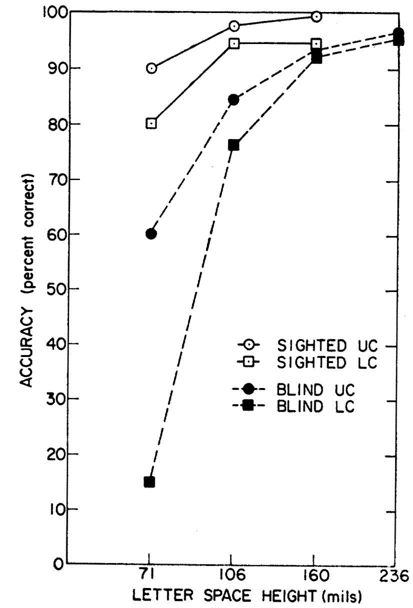

Fig. 9. Reading aid output legibility as a function of letterspace height. Recognition accuracy on random strings of upper case letters and numbers (UC), and on lower case letters (LC), was measured for sighted subjects observing the light display and for blind subjects using the stimulator display.

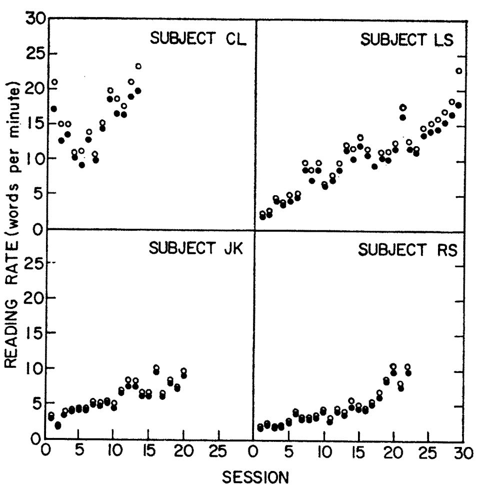

Fig. 10. Reading performance of four subjects with the 24 × 6 reading aid. The open circles indicate the average line reading rate, not including the time required to move the probe from the end of one line to the beginning of the next. The solid points indicate the rates obtained when this line change time is included. Subject JK was unable to continue after the twentieth session.

Two measures of reading rate are given in Fig. 10. The higher rate does not include the time required to move the probe from the end of a line of type to the beginning of the next line, and to accurately position it, while the lower rate does. The difference in these two rates is almost 20 percent in the case of subject CL.

The importance of this line change time in limiting reading rate is illustrated in Fig. 11. The curves in this figure are based on a 12-word line. At their present level of performance, the subjects require from two to ten seconds to change lines. If their continuous line reading rate is 25 words/min, this line change time reduces the overall reading rate to about 19–23 words/min, and if the continuous line rate is 50 words/min, then the overall

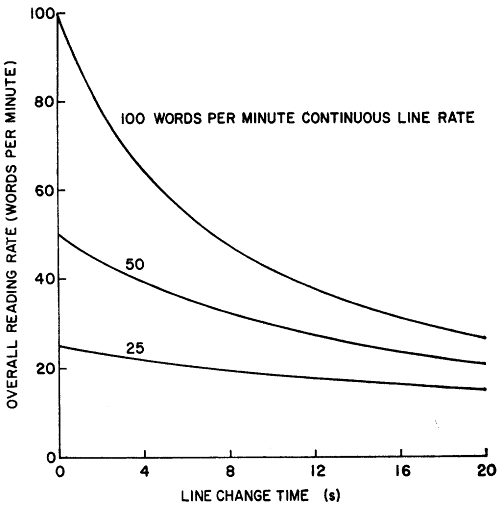

Fig. 11. Effect of line change time on overall reading rate. These curves are calculated based on a 12-word line.

limit is 30–43 words/min. A larger array and a smaller and lighter probe would undoubtedly improve this situation.

There are two ways in which reading with this device is more difficult than with the computer stimulation with which we obtained rates of over 30 words/min. First, with this device the subjects must manually track the line of type with considerable precision, and this requires concentration. Secondly, the effects of sampling and quantization are apparent in the output display, while the computer simulation produced perfect letters.

Future Possibilities

While we believe our present reading aid performs adequately for useful reading, there are many areas in which improvements could be made. This aid was constructed to test the design concepts, and, therefore, readily available components were used rather than optimum ones. Thus, if a number of such devices were to be built, several changes in construction would be appropriate. For example, if an integrated array of phototransistors of the appropriate size and sensitivity were available, the construction would be significantly simplified since the fiber optics could be eliminated, and the amount of handling of discrete components could be greatly reduced. Performance would also be increased since a great deal of the channel-to-channel variability is due to fiber transmission and fiber–transistor interface variations. Also, about half the light is lost in the fibers. If the integrated photocell array was small enough to permit optical demagnification instead of magnification, the hand-held probe could be significantly smaller. This should reduce the line-change time and, correspondingly, increase the overall reading rate.

Another innovation to reduce manufacturing costs would be to multiplex the channels rather than have completely parallel circuitry. This would also increase performance since greater amplification could be afforded on the few parallel channels remaining and since fewer wires would connect to the optical probe, thus making manipulation of the probe easier. While the present overall size of the reading aid is manageable, a size reduction could be easily obtained and would result in much greater convenience. Multiplex electronics for this purpose have been built in our laboratories and will be incorporated in future models.

Enhanced performance with the reading aid could also be expected, with little increase in device complexity, by preserving some of the analog information from the photosensors and displaying several shades of tactile "gray." Also, some of the character "clean-up" techniques, commonly used in automatic character recognition equipment, may improve the images obtainable with this reading aid.

More significant extensions of the present design involve attempts to greatly increase the possible reading rate. One approach to this goal is to increase the field of view greatly so that "word-at-a-time" reading could be obtained. However, if the same resolution is maintained over the entire field of view, the number of channels would become very high. It might be possible to reduce the number of channels needed by using a variable resolution across the field of view.

Another approach presently being considered is to provide an auditory output, in addition to the tactile display, which consists of a combination of spoken words and spoken letters. This output could be achieved by using telephone lines to transmit partially processed photosensor signals to a central time-shared computer system. The central facilities could then recognize the letters and produce the spoken output. In this case the primary utility of the tactile display would be to aid in tracking, and as a back-up for patterns that are not machine recognizable. If such a system could permit reading rates of several hundred words/min, its value would be established. Fortunately, the reading rate possible with such a system can be measured through computer simulation before such a system is developed. We are presently experimenting with such a simulation.

The system described here can also be adapted for viewing the environment by appropriate changes in the optics. We are also presently exploring this possibility.

Acknowledgment

The author would like to express his deep appreciation to Prof. J. G. Linvill of Stanford University for his invaluable council and continual encouragement. The technical contributions and generous assistance of the students associated with this project, especially J. Taenzer, C. H. Rogers, R. C. Joy, and J. W. Hill, are gratefully acknowledged.

J. Baer and J. Gill of Stanford Research Institute were mainly responsible for the design and construction of the reading aid model described here. Miss Carolyn Weihl conducted most of the tests involving subjects. Dr. L. D. Harmon, of the Bell Telephone Laboratories, kindly read the manuscript and made many helpful suggestions.

References

[1] E. E. Fournier d'Albe, "On a type-reading optophone," *Proc. Roy. Soc.* (London), vol. 90, no. A619, pp. 373–375, 1914.

[2] V. K. Zworykin and L. E. Flory, "An electronic reading aid for the blind," *Proc. Am. Phil. Soc.*, vol. 91, no. 2, pp. 139–142, 1947.

[3] J. S. Abma, "The battelle aural reading device for the blind," in *Human Factors in Technology*, E. Bennett, J. Degan, and J. Spiegel, Eds. New York: McGraw-Hill, 1963, pp. 315–325.

[4] G. C. Smith, "The development of recognition and direct translation reading machines for the blind," *Proc. Internat'l Conf. on Sensory Devices for the Blind*, R. Dufton, Ed. Bristol: J. W. Arrowsmith, 1967, pp. 367–388.

[5] H. Freiberger and E. F. Murphy, "Reading machines for the blind," *IRE Trans. Human Factors in Electronics*, vol. HFE-2, pp. 8–19, March 1961.

[6] J. G. Linvill and J. C. Bliss, "A direct translation reading aid for the blind," *Proc. IEEE*, vol. 54, pp. 40–50, January 1966.

[7] F. W. Campbell, J. J. Kulikowski, and J. Levinson, "The effect of orientation on the visual resolution of gratings," *J. Physiol.*, vol. 187, pp. 427–436, 1966.

[8] J. L. Harris, "Image processing as it relates to the human system," in *Current Developments in Optics and Vision*. Washington, D.C.: NAS-NRC, 1968, pp. 78–88.

[9] J. Rabinow, "The present state of the art of reading machines," in *Pattern Recognition*, L. N. Kanal, Ed. Washington, D.C.: Thompson Book Co., 1968, pp. 3–30.

[10] H. Kazmierczak, "Stage of development of automatic character recognition and complex reading machines for the blind in Europe," *Proc. Internat'l Congress on Technology and Blindness*, vol. 1. New York: American Foundation for the Blind, 1963, pp. 261–278.

[11] J. W. Hill and J. C. Bliss, "Modeling a tactile sensory register," *Perception and Psychophys.*, vol. 4, pp. 91–101, August 1968.

[12] D. E. Troxel, "Experiments in tactile and visual reading," *IEEE Trans. Human Factors in Electronics*, vol. HFE-8, pp. 261–263, December 1967.

[13] R. C. Joy and J. G. Linvill, "Phototransistor operation in the charge storage mode," *IEEE Trans. Electron Devices*, vol. ED-15, pp. 237–248, April 1968.

[14] F. Geldard, "Pattern perception by the skin," in *The Skin Senses*, D. R. Kenshalo, Ed. Springfield, Ill.: Charles C. Thomas, 1968, pp. 304–319.

[15] A. Spiridon, "Vibro-tactile pattern perception," Sc.D. dissertation, Dept. Elec. Engrg, Massachusetts Institute of Technology, Cambridge, 1966.

[16] J. W. Hill and J. C. Bliss, "Perception of sequentially presented tactile-point stimuli," *Perception and Psychophys.* vol. 4, pp. 289–295, November 1968.

[17] S. Weinstein, "Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality," in *The Skin Senses*, D. R. Kenshalo, Ed. Springfield, Ill.: Charles C. Thomas, 1968, pp. 195–219.

[18] K. C. Mukherjee, "The duration of cutaneous sensation (I) and the improvement of its sensible discrimination by practice." *J. Expt. Psych.*, vol. 16, pp. 339–342, 1933.

[19] M. J. T. Fitzgerald, "The innervation of the epidermis," in *The Skin Senses*, D. R. Kenshalo, Ed. Springfield, Ill.: Charles C. Thomas, 1968, pp. 61–81.

[20] W. H. Talbot, I. Darian-Smith, H. H. Kornhuber, and V. B. Mountcastle, "The sense of flutter-vibration: comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand," *J. Neurophysiol.*, vol. 31, no. 2, pp. 301–333, 1968.

[21] V. B. Mountcastle, W. H. Talbot, I. Darian-Smith, and H. H. Kornhuber, "Neural basis of the sense of flutter-vibration," *Science*, vol. 155, pp. 597–600, February 3, 1967.

[22] U. Lindblom, "Properties of touch receptors in distal glabrous skin of the monkey," *J. Neurophysiol.*, vol. 28, pp. 966–985, Sptember 1965.

[23] U. Lindblom and L. Lund, "The discharge from vibration sensitive receptors in the monkey foot," *Expt. Neurol.*, vol. 15, pp. 401–417, 1966.

[24] R. B. Arps, R. L. Erdmann, A. S. Neal, and C. E. Schlaepfer, "Optimal resolution for image processing of printed matter," *1968 IEEE Internat'l Conf. on Communication Record*, vol. 4, pp. 838–843, 1968.

The Human as an Optimal Controller and Information Processor

SHELDON BARON, MEMBER, IEEE, AND DAVID L. KLEINMAN, MEMBER, IEEE

Abstract—A mathematical model of the instrument-monitoring behavior of the human operator is developed. The model is based on the assumption that the operator behaves as an optimal controller and information processor, subject to his inherent physical limitations. The resulting model depends explicitly on the control task and the control actions. Provision is made for the ability to obtain information from the peripheral visual field. There are no restrictions on signal coupling.

The specific characteristics of the operator's visual sampling behavior are predicted by solving a nonlinear, deterministic optimization problem. A two-axis compensatory tracking example is investigated, and the results exhibit the general characteristics expected of a human operator performing a similar task.

Manuscript received April 17, 1968; revised November 14, 1968. This work was supported by the National Aeronautics and Space Administration under Contract NAS12-104.

The authors are with Bolt Beranek and Newman, Inc., Cambridge, Mass. 02138.

I. Introduction

WE ARE currently investigating the problem of manual control in complex multivariable situations with the aim of obtaining a more complete mathematical description of the human operator's control and information-processing behavior. The model we are studying is rooted in modern control theory. The main assumption underlying its development is that the well-trained well-motivated operator behaves in a near-optimal manner, subject to certain inherent constraints. The model contains elements for describing the operator's instrument-monitoring, data-reconstruction, and control behavior, as well as means for representing his inherent limitations. Thus, the operator's control behavior is assumed to be that of an ideal feedback con-